Cybersecurity Report

30.12.2025 – 8.1.2026

Stories

OpenAI potvrdzuje, že ChatGPT nebude používať zdravotné údaje na tréning modelov

OpenAI oznámilo, že zdravotné informácie používateľov nebudú využité na tréning ChatGPT modelov, čím chce zmierniť obavy týkajúce sa ochrany súkromia a citlivých dát. Spoločnosť uviedla, že dáta z lekárskych správ, výsledkov testov alebo iných zdravotných záznamov zaslané do ChatGPT sa nebudú pridávať do tréningových datasetov ich jazykových modelov a nebudú slúžiť na zlepšovanie generatívnej AI. Informácia prichádza v kontexte rastúcich diskusií o tom, ako sú citlivé dáta spracovávané v AI službách a aké kontroly by mali existovať na ich ochranu. OpenAI tak explicitne oddelila používanie zdravotných údajov od procesu učenia modelov, hoci si používatelia stále musia byť vedomí všeobecných obmedzení ochrany súkromia pri zdieľaní akýchkoľvek osobných informácií s AI nástrojmi.

https://www.bleepingcomputer.com/news/artificial-intelligence/openai-says-chatgpt-wont-use-your-health-information-to-train-its-models/

Kritická RCE zraniteľnosť (CVSS 10.0) ohrozuje platformu n8n a stovky tisíc inštancií

Open-source workflow automatizačná platforma n8n varuje pred kritickou zraniteľnosťou (CVE-2026-21858) s CVSS skóre 10.0, ktorá umožňuje neautentifikovaným útočníkom spustiť ľubovoľný kód na zraniteľných inštanciách, získať prístup k súborom, tokenom a potenciálne prevziať ich úplnú kontrolu. Táto chyba, prezývaná „Ni8mare“, vyplýva z nesprávnej validácie vstupov vo webhookoch a spracovaní obsahu, čo útočník môže využiť bez potreby prihlasovacích údajov a dosiahnuť full system compromise. Odhady ukazujú, že desiatky tisíc internetovo vystavených n8n serverov zostávajú zraniteľných, pričom opravná verzia 1.121.0 alebo novšia bola vydaná v novembri 2025 a užívateľom sa dôrazne odporúča okamžitá aktualizácia. Okrem toho bola riešená aj ďalšia kritická RCE zraniteľnosť (CVE-2026-21877) v Git node, ktorá umožňuje autentifikovaným útočníkom spúšťať neoverený kód a kompletne kompromitovať n8n inštancie.

https://thehackernews.com/2026/01/n8n-warns-of-cvss-100-rce-vulnerability.html

---

Prompt injection zneužíva funkciu „Memory“ v ChatGPT na podvodné odpovede

Výskumníci upozorňujú na ďalší vektor zneužitia veľkých jazykových modelov súvisiaci s funkciou Memory v ChatGPT, ktorý môže viesť k *prompt injection* útokom so skrytými škodlivými inštrukciami. Funkcia Memory, ktorá si pamätá kontext a predchádzajúce interakcie používateľa, môže byť zneužitá tak, že útočník schová nežiaduce príkazy do dlhodobo uložených poznámok alebo do repetitívnych interakcií. Keď model neskôr reaguje na legitímnu požiadavku, môže *implicitne* vykonať tieto ukryté prompt-injekcie, čo vedie k nečakanému a bezpečnostne nevhodnému spracovaniu textu. Útoky tohto typu obchádzajú bežné „input sanitization“ mechanizmy, pretože škodlivé inštrukcie sa nenachádzajú priamo v aktuálnom prompt-e, ale v pamäti modelu.

https://www.darkreading.com/endpoint-security/chatgpt-memory-feature-prompt-injection

---

Zoom Stealer: škodlivé prehliadačové rozšírenia zbierajú dáta z firemných videohovorov

Výskumníci odhalili kampaň škodlivých prehliadačových rozšírení zameranú na získavanie citlivých informácií z Zoom videohovorov a súvisiacich pracovných dát. Tieto rozšírenia, ktoré sa maskujú ako legitímne nástroje na zlepšenie používateľského zážitku, po inštalácii sledujú aktivitu v prehliadači, zachytávajú obsah Zoom schôdzok, identifikátory meetingov, tokeny relácií a ďalšie dáta súvisiace s firemnými stretnutiami. Útočníci ich následne odosielajú na backend server, kde ich môžu zneužiť na neoprávnený prístup k účtom, replay útoky či šírenie phishingu vo firemnom prostredí. Kampaň zdôrazňuje riziko používania neoverených rozšírení v pracovnom prehliadači.

https://www.bleepingcomputer.com/news/security/zoom-stealer-browser-extensions-harvest-corporate-meeting-intelligence/

Google najíma „AI quality“ inžinierov po náraste hallucinácií v Search AI

Google reaguje na pretrvávajúce AI hallucinations v Google Search generovaných odpovediach zvýšením zamestnancov zameraných výlučne na kvalitu AI odpovedí. Tieto chyby, pri ktorých generované výsledky môžu byť nepresné alebo zavádzajúce, vedú k zlej používateľskej skúsenosti a potenciálne riskantným situáciám, keď sa používatelia spoliehajú na automaticky generovaný obsah pre rozhodovanie. Spoločnosť teda rozširuje tím „AI Answers Quality Engineers“, ktorí majú monitorovať, testovať a opravovať nesprávne alebo nebezpečné odpovede v reakciách generovaných AI v rámci vyhľadávania. Google zároveň upravuje interné metriky a procesy, aby zlepšil presnosť výstupov a znížil výskyt halucinácií, pričom dôraz kladie na kombináciu automatizovaných testov a ľudskej spätnej väzby. Táto iniciatíva je súčasťou širšieho úsilia zlepšiť dôveryhodnosť a spoľahlivosť AI funkcií v produktoch Google.

https://www.bleepingcomputer.com/news/microsoft/microsoft-exchange-online-outage-blocks-access-to-mailboxes-via-imap4/